中图法分类号:TP18;TP391　　文献标识码:A　　文章编号:1006-8961(2026)10-3441-19

论文引用格式:Li D X, Xie W Y and Li Y S. 2026. Embodied intelligence methods for deep space exploration: a survey. Journal of Image and Graphics, 31(10):3441-3459(李岱勋, 谢卫莹, 李云松. 2026. 面向深空探测的具身智能方法研究进展. 中国图象图形学报, 31(10):3441-3459)

[DOI:10.11834/jig.260123]

面向深空探测的具身智能方法研究进展

李岱勋,谢卫莹*,李云松

西安电子科技大学通信工程学院,西安　710071

摘　要: 随着深空探测任务不断向高自主、长时序和复杂环境作业发展,传统依赖地面的遥控遥操任务模式已难以满足月球、火星及更远深空环境下实时、可靠和高效作业的需求。面向这一发展趋势,具身智能逐渐成为深空探测的重要研究方向,其核心在于推动探测系统由单一指令执行向感知、决策与执行一体化自主运行演进,从而提升深空任务中的环境感知能力与自主决策水平。本文系统回顾了深空探测智能化方法的研究进展,围绕深空感知能力受限、自主决策能力不足等关键问题,梳理了深空视觉感知方法与深空具身智能方法的研究现状,并对相关数据集、典型任务及评价方式进行了归纳总结。在此基础上,进一步从感知泛化、可信决策、系统协同与轻量化部署等方面,对深空探测具身智能的发展方向进行了展望,旨在为深空探测自主智能技术的研究与应用提供新的技术路径。

关键词: 深空探测;具身智能;视觉感知;火星探测;月球探测;深度学习

Embodied intelligence methods for deep space exploration: a survey

Li Daixun, Xie Weiying*, Li Yunsong

College of Communication Engineering, Xidian University, Xi'an　710071, China

Abstract: As deep space exploration missions continue to advance toward higher autonomy, longer operational duration, and more complex environments, the traditional task paradigm dominated by ground-based teleoperation can no longer satisfy the requirements for real-time, reliable, and efficient operations in lunar, Martian, and farther deep-space scenarios. Under conditions such as significant communication delay, limited prior environmental knowledge, strict onboard resource constraints, and high trial-and-error cost, future deep space platforms are expected to move beyond single-command execution toward integrated autonomous operation. In this context, embodied intelligence is gradually emerging as an important research direction for deep space exploration. The reason is that it provides a unified framework for coupling environmental perception, autonomous decision-making, and physical execution, which improves the adaptability and autonomy of exploration systems in long-duration extraterrestrial missions. This study systematically reviews recent progress in intelligent methods for deep space exploration from the perspective of embodied intelligence. Focusing on two central challenges—limited perception capability in highly unstructured extraterrestrial environments and insufficient autonomous decision-making capability under long-horizon, high-risk, and resource-constrained conditions—the study reviews recent advances in deep space visual perception and deep space embodied intelligence, and summarizes representative datasets, typical tasks, and

收稿日期:2026-03-08;修回日期:2026-05-07;预印本日期:2026-05-14

*通信作者:谢卫莹　wyxie@xidian.edu.cn

基金项目:国家自然科学基金项目(62531018,U24B20136);教育部学科先导突破项目(JYB2025XDXM105)

Supported by: National Natural Science Foundation of China (62531018, U24B20136); Fundamental and Interdisciplinary Disciplines Breakthrough Plan of the Ministry of Education of China (JYB2025XDXM105)

evaluation protocols. For visual perception, the survey first reviews traditional image processing-based methods, including stereo terrain reconstruction, visual odometry, photogrammetric modeling, and large-scale localization. These methods established the technical foundation for early planetary exploration because they are grounded in explicit geometric constraints and therefore offer strong interpretability and engineering reliability. However, their effectiveness is increasingly challenged by weak textures, repeated patterns, severe illumination changes, occlusion, and long-range cumulative drift in lunar and Martian environments. The study then examines deep learning-based visual perception methods, which cover terrain semantic classification and segmentation, cross-domain adaptation, and multimodal terrain understanding. Compared with traditional geometry-driven approaches, deep models provide stronger semantic expressiveness and improved environmental understanding in complex planetary scenes. At the same time, their performance still depends heavily on data coverage, annotation quality, training strategy, and sim-to-real consistency, while onboard deployment is constrained by limited computation and energy budgets. From this perspective, the development of deep space perception is gradually shifting from closed-dataset recognition accuracy toward robust representation, long-term adaptability, and task-oriented environmental understanding. From the perspective of embodied intelligence, the study categorizes deep space tasks into three major classes: autonomous visual navigation, autonomous mission planning, and autonomous scientific operation. On this basis, the methodological evolution of deep space embodied intelligence can be viewed in three stages. The first stage is represented by teleoperation-centered control models, in which planning and key decisions are mainly completed on the ground, while the onboard system executes predefined instructions with limited rule-based local autonomy. Although this paradigm has been essential for ensuring mission safety and controllability, its strong dependence on human intervention limits its capability to support future long-duration missions in uncertain environments. The second stage is characterized by deep-learning-enhanced local autonomy. In this stage, learning models are embedded into specific modules of conventional engineering pipelines to improve terrain understanding, candidate path evaluation, risk estimation, target selection, and onboard data screening, while the overall decision and control chain remains largely interpretable and verifiable. The third stage involves multimodal large models and vision-language-action systems, which introduce higher-level cognitive capability into deep space systems. By unifying multimodal observations, state information, and task instructions in a shared semantic space, these methods enable cross-modal reasoning, task decomposition, dynamic replanning, and long-horizon decision organization. The survey argues that such models are promising for future high-autonomy missions. However, under current engineering conditions, they are more suitable as high-level cognition and task-orchestration modules than as direct substitutes for low-level planning and control. The study further reviews datasets and evaluation methods for visual perception and embodied intelligence. Compared with terrestrial benchmarks, publicly available deep space datasets remain limited in scale, annotation richness, and task diversity. For embodied intelligence, the challenge is even more fundamental, considering that datasets must support temporality, interaction, and closed-loop evaluation by including not only observations but also pose, action, state, and feedback information. As a result, high-fidelity simulation environments and digital twin-style benchmarks currently play an important role in supporting reproducible system evaluation. Correspondingly, evaluation protocols are also evolving from static recognition metrics toward task-level criteria involving safety, effectiveness, efficiency, robustness, and long-horizon reliability. Based on the abovementioned survey, the study discusses several future directions for deep space embodied intelligence, including perceptual generalization in strongly unstructured environments, trustworthy autonomous decision-making, system-level collaboration across perception and control modules, and lightweight deployment on resource-constrained platforms. In summary, deep space embodied intelligence is emerging as a system-level intelligence problem shaped by extreme extraterrestrial environments, long-duration mission chains, strict resource limitations, and high reliability requirements. Its future advancement depends not only on improving individual algorithms but also on establishing an integrated framework that unifies perception, decision-making, execution, validation, and deployment for real deep-space missions. This survey aims to provide a structured reference and clearer analytical perspective for the further development and application of autonomous intelligence technologies in deep space exploration.

Key words: deep space exploration; embodied intelligence; visual perception; Mars exploration; lunar exploration; deep learning

0 引 言

随着深空探测活动的推进，人类对月球、火星及更远天体的探测正由传统的“地面规划与遥控遥操为主”的任务模式（李立春 等，2022），逐步迈向高智能、高可信以及高自主的无人系统发展阶段（刘继忠等，2024）。近年来，国际月球科研站基本型建设（张哲 等，2022）、火星巡视探测（Whetsel 等，2025）、小行星采样返回以及行星际穿越探测等重大任务不断推进，对探测器长期自主运行能力提出更高要求。在通信时延大、环境复杂未知以及计算资源受限的深空环境下，传统依赖地面控制的任务执行模式已难以满足复杂任务对实时性与可靠性的要求。研究表明，未来行星探测任务需要更高水平的自主能力，以支持探测器在未知环境中完成自主导航、任务规划与科学作业等任务，从而显著提升任务效率与科学产出。在这一背景下，具备环境感知、决策规划与动作执行能力的具身智能系统逐渐成为深空探测技术发展的重要方向。

从任务形态上看，深空探测任务呈现出明显的多样性与层级性特征。一方面，深空探测任务通常涵盖从近地环绕到深空巡航的多阶段过程（宁晓琳等，2023），包括轨道段的持续观测与资源调度、着陆段及地表段的巡视、导航、采样与原位探测等多类型任务。这些任务不仅具有时间跨度长、环境不确定性高等特点，还往往涉及多种传感器协同和复杂作业流程。另一方面，探测器平台类型也在不断丰富，包括近地轨道器、地表作业着陆器以及地面探测器等（叶培建 等，2019）。其中，地面探测器进一步发展为多种形态，如轮式或履带式探测车、跳跃式或飞跃式月表探测平台，以及具备局部机动能力的火星探测无人机等新型平台。不同类型探测器在运动方式、作业模式、传感器配置以及计算资源方面存在显著差异，这使得深空探测对具身智能方法在环境适应性、任务通用性以及工程可实现性方面提出了更高的要求。

与此同时，深空探测场景通常表现出典型的非结构化环境特征。以月球地表为例，其地貌由陨石坑、岩块、坡地与松散覆盖物等复杂结构构成，空间形态高度不规则，并受到光照剧烈变化、遮挡效应以及传感器视角受限等因素的影响（邢琰 等，2024）。这类环境通常呈现高度非结构化特征，且先验信息有限、数据采集能力受限、试错成本高昂，使得深空探测任务在自主运行过程中面临较大的不确定性风险。此外，深空探测系统还需要在严格的资源约束条件下长期稳定运行，包括计算能力、能耗预算以及通信带宽等方面的限制（唐玉华 等，2025）。与通用机器人或自动驾驶等应用场景相比，深空探测在任务执行、数据获取与系统验证方面具有更强的工程约束，因此相关智能方法不仅需要具备环境适应能力，还需要兼顾可靠性、可控性以及长期运行稳定性。

近年来，围绕深空探测相关问题，已有学者从不同角度对任务需求与关键技术进行了较为系统的综述。例如，一些研究对行星探测任务规划与执行流程进行了总结，分析了不同深空任务在导航、通信与资源调度方面的共性问题（傅晓晶 等，2023）；部分工作聚焦于深空探测中的视觉感知与遥感解译方法，对传统图像处理与深度学习技术在行星表面理解中的应用进行了归纳；也有综述针对深空机器人、行星车控制或星载智能系统展开讨论（金文马 等，2024），从控制模型、任务分解与系统架构等层面总结了相关研究进展。然而，从整体上看，面向深空探测的具身智能方法仍缺乏系统性的梳理与统一视角的综述。现有工作往往分别从视觉感知、控制方法或单一任务场景展开讨论，尚未从具身智能系统闭环的角度，对深空探测具身智能方法进行系统总结。特别是在非结构化复杂环境以及高自主运行需求并存的背景下，不同研究方向之间的内在联系与技术演进脉络仍有待进一步梳理。因此，有必要从具身智能系统整体架构出发，对深空探测中的关键方法与技术路线进行系统归纳，以更清晰地揭示该领域的发展趋势与未来研究方向。

本文从具身智能系统整体架构出发，对深空探测具身智能方法进行系统综述。具体而言，本文首先分析深空探测具身智能系统面临的关键挑战，包括感知能力受限、自主决策能力不足的问题；随后从视觉感知方法、具身智能决策方法两个方面，对近年来相关研究进展进行系统归纳与对比分析，并总结典型方法的发展脉络与技术特点；最后结合深空探测任务的发展趋势，对未来深空具身智能系统在多模态感知、自主任务规划以及高效边缘部署等方面的潜在研究方向进行讨论。总结而言，本文旨在为

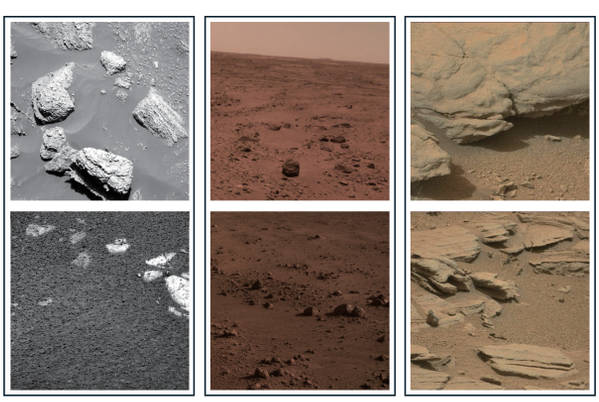
深空探测具身智能方法的研究提供系统化参考，并为后续深空自主探测技术的发展提供理论与方法层面的借鉴。

1 深空具身智能方法的挑战

深空探测任务通常面向月球、火星及更远行星体的长期自主运行需求，涵盖环境感知、导航规划与原位作业等复杂任务形态。受探测距离远、通信时延大以及环境先验信息匮乏等因素影响，探测器在实际运行过程中难以持续依赖地面遥控遥操，因而需要在边端平台具备一定程度的自主运行能力。在此背景下，如何在深空非结构化环境和受限平台条件下构建具备工程可行性的具身智能系统，成为当前深空探测研究中亟需关注的问题，具体挑战描述如下。

1.1 深空感知能力受限

深空探测任务通常运行于高度非结构化的自然环境中，其感知条件与地面应用场景存在显著差异。一方面，月球、火星等天体表面缺乏人工基础设施和稳定参照物，如图1所示，深空地形结构复杂且空间分布高度不规则，感兴趣目标与背景之间的区分度有限，给环境理解带来先天障碍。另一方面，深空探测器在作业过程中常受到极端光照变化、遮挡效应以及传感器姿态和视角受限等因素影响，使得获取的观测信息存在噪声大、不完整和不连续等问题。



(a) 传感器噪声大　(b) 空间分布不规则　(c) 地形结构复杂

图1　深空传感器数据挑战

Fig. 1　Deep space sensor data challenge

((a) high sensor noise; (b) irregular spatial distribution; (c) complex terrain structure)

从数据条件来看，深空探测的感知能力还受到先验知识与验证机制的制约。与地面视觉任务依赖大规模、可重复采集和持续标注的数据集不同，深空任务中可获取的真实数据数量有限，且数据分布往往与地面模拟或实验环境存在显著偏差。这种少样本、强分布差异的特征，使得感知模型在训练阶段难以充分覆盖复杂场景变化，在实际任务中容易出现表征不足或泛化能力下降的问题。此外，深空探测对感知结果的可靠性与稳定性要求远高于多数地面应用。感知误差不仅影响局部环境理解，还可能在后续决策与执行过程中被逐步放大，从而引发路径规划失败或作业风险累积。因此，如何在受限观测条件下获得具有工程可用性的环境表征，是深空具身智能面临的首要挑战。

1.2 自主决策能力不足

在深空探测任务中，具身智能系统不仅需要理解环境状态，还需要基于感知信息进行长期自主决策。典型任务包括自主视觉导航、自主任务规划以及复杂作业控制等，这些任务往往具有长时间跨度、多阶段目标和严格安全约束等特点。然而，现有自主决策方法在深空非结构化环境中的适应能力仍然有限。传统基于模型或规则的决策与控制方法在可解释性和稳定性方面具有一定优势，但通常依赖较为完备的环境建模与先验假设，在面对未知地形、动态风险和复杂任务约束时灵活性不足。与此同时，基于学习的方法在感知与决策耦合建模方面展现出潜力，但其有效性往往依赖大量交互数据和反复试错过程，这与深空探测高代价、低容错的工程实际目前尚存在一定矛盾。

此外，深空任务中的科学作业通常表现为强耦合的“感知—决策—执行”闭环过程。感知结果的不确定性、推理决策的安全风险以及执行过程的动态扰动，共同增加了系统建模的复杂度，使其难以通过简单的短时策略完成深空探测自主作业，这在一定程度上制约了无人系统的智能化自主运行能力。因此，如何在复杂约束条件下实现感知、推理与执行链路的协同闭环，已成为提升深空探测系统自主能力的关键问题。总体来看，上述问题共同构成了深空探测具身智能方法研究的现实约束。后续内容将围绕这些挑战，从视觉感知、具身智能决策两个方面，对相关研究进展进行系统梳理与分析。

2 深空探测视觉感知方法

2.1 基于传统图像处理的视觉感知方法

视觉感知是深空自主导航与定位的基础，其核心任务包括利用载体搭载的视觉传感器获取环境信息，通过图像几何与匹配推导三维地形结构，并在此基础上估计探测器的空间位置。传统图像处理方法在这一领域的研究可以追溯到早期行星车任务，形成了一条从地形建模、立体几何测量、视觉里程计和全局定位融合的方法链。

早期地形建模工作的代表性成果是针对火星表面三维地形的多分辨率重建。Olson 等人（2007）提出了基于宽基线立体匹配的多尺度地形融合方法，通过融合着陆器、轨道与车载表面图像建立大范围三维地图，为漫游车导航和定位提供了多层次地形表示。这一工作为后续基于单目/双目视觉的三维重建和导航算法奠定了基础。在此基础上，立体视觉作为深空定位与导航的核心技术之一，广泛应用于行星车任务。马友青等人（2014）针对月球车运动环境，提出了基于立体图像的导航与定位算法，通过空间三角测量从图像对中恢复三维场景深度信息并进一步估计探测器姿态。类似地，Li 等人（2013）基于立体视觉构建视觉里程计（visual odometry，VO）系统，用于估计火星车在不平坦沙地环境中的实时运动状态，通过特征匹配和几何重建实现位置变化的有效估计。

更具代表性且工程上影响深远的是美国国家航空航天局（National Aeronautics and Space Administration，NASA）MER（Mars exploration rovers）任务中视觉里程计的系统实现。Maimone 等人（2007）利用 MER 任务的立体导航摄像机对地面特征进行跟踪，设计了特征基视觉里程计算法，使探测器能够在火星真实环境下实现位置估计与滑移补偿，显著提高了导航安全性和里程精度。为提升视觉定位的准确度与鲁棒性，研究者进一步将视觉里程计与几何重建方法相结合（耿迅和徐青，2022），实现更大范围的定位与地形理解。例如，通过在视觉里程计和光束法平差之间建立联动机制，能够在短时视觉跟踪与全局地图一致性之间取得平衡，从而提高长航程定位精度。此外，有工作尝试将车载视觉信息与轨道遥感数据进行跨视角匹配（Li 等，2025b），实现绝对定位和累积误差矫正，这对于深空任务中长距离、长期导航尤为重要。

通过立体视觉重建、视觉里程计以及跨尺度图像配准等技术，传统视觉方法逐步实现了从局部地形感知到较大范围导航定位的能力扩展。从技术范式上看，这类方法本质上建立在几何约束与特征匹配之上，通过显式成像模型恢复三维结构与位姿信息，因此具有较强的可解释性和较高的工程成熟度。在早期深空探测任务中，这一范式能够较好地适应结构相对稳定、纹理较为清晰的场景需求，是深空自主导航与定位的重要技术基础。尤其是在计算资源受限、模型行为需可验证的工程背景下，传统方法在可靠性和可控性方面仍具有不可替代的价值。

然而，随着探测任务逐步转向更复杂的非结构化环境，传统方法的局限性也愈发明显。一方面，月球和火星地表普遍存在弱纹理、重复纹理、光照剧烈变化以及局部遮挡等问题，导致基于特征匹配的几何方法稳定性下降；另一方面，视觉里程计等局部估计方法难以从机制上抑制长时序误差累积，在长航程任务中容易出现定位漂移。因此可以认为，传统视觉方法在深空场景中的优势主要体现在可解释性与工程可验证性，而其瓶颈则集中表现为对复杂环境适应能力不足、对高层语义信息利用有限以及对长期自主任务支撑能力不足。这些结构性限制也直接推动了后续研究向数据驱动的深度学习感知方法演进。

2.2 基于深度学习的视觉感知方法

随着深度学习技术在计算机视觉领域的快速发展，如表 1 所示，研究者开始探索基于深度学习的深空探测视觉感知系统在复杂环境下的表达能力与鲁棒性。相较于基于传统图像处理的几何建模与特征匹配方法，深度学习方法通过端到端学习方式自动提取高层语义特征，在非结构化环境理解、地形语义建模以及复杂视觉模式识别等方面展现出更强的潜力，逐渐成为深空视觉感知研究的重要方向。

2.2.1 基于深度学习的地形语义分类与分割方法

地形语义理解是深空探测视觉感知中的核心任务之一，其目标在于从复杂地表图像中区分不同地形类型，为导航规划与作业决策提供高层环境信息。Rothrock 等人（2016）较早将卷积神经网络引入火星地形分类任务，提出了基于深度学习的地形语义预测框架，显著提升探测车对可通行与危险地形的识

表1　深空探测视觉感知方法归纳

Table 1　Overview of visual perception methods for deep space exploration

任务类别	文献	方法类别	数据来源	主要任务
导航定位	Olson等人(2007)	立体视觉	火星探测器/轨道飞行器影像	火星三维地形建模
	Maimone等人(2007)	视觉里程计	火星探测器导航相机	火星车视觉里程计
	Li等人(2013)	立体视觉	月球探测车	月球车里程计
	马友青等人(2014)	立体导航	月球探测车	月球车导航定位
	耿迅和徐青(2022)	摄影测量	火星车高分辨率立体相机/月球勘测轨道飞行器窄角相机	行星摄影测量
	Driver等人(2023)	深度学习	小行星	自主导航
	Chiuchiarelli等人(2024)	深度学习	火星探测车	自主导航
目标检测	Bickel等人(2021)	深度学习	火星勘测轨道飞行器/月球勘测轨道飞行器	落石检测
	Lyu等人(2023)	深度学习	“祝融号”火星车	岩石检测
	曹学欢等人(2025)	深度学习	“祝融号”火星车	岩石检测
分类分割	Rothrock等人(2016)	深度学习	火星2020火星车	地形分类
	Swan等人(2021)	深度学习	“好奇号”、“机遇号”和“勇气号”火星探测车	地形分类
	Barrett等人(2022)	深度学习	火星探测车	地貌分类
	Vincent等人(2022)	深度学习	“好奇号”、“毅力号”火星探测车	地形分类
	Goh等人(2022)	深度学习	火星探测车	地形分类
	Tan等人(2024)	深度学习	火星探测车	地形分类
	Mohammad等人(2024)	深度学习	火星探测车	地形分类
	Castilla-Arquillo等人(2025)	深度学习	火星探测车	地形分类
	Panambur和Parente(2025)	深度学习	火星探测车	地形分类
	Tan等人(2025)	深度学习	火星探测车	零样本地形分类
	Jiao等人(2025)	深度学习	月球探测车	月球障碍检测

别能力。该工作验证了深度学习在行星地表图像理解中的可行性,为后续研究奠定了基础。

针对深空探测数据标注成本高、样本规模有限的问题,Goh等人(2022)提出了基于半监督学习的火星地形分割方法,通过引入未标注数据参数与训练,有效缓解了数据稀缺对模型性能的制约。进一步地,Mohammad等人(2024)系统评估了U-Net、DeepLabV3+等主流语义分割网络在火星地表图像上的表现,并结合数据增强策略提升了模型对复杂地形边界的刻画能力。

随着网络结构的发展,一些研究开始引入Transformer等新型深度模型以增强对空间上下文关系的建模能力。例如,Ma等人(2025)通过将深度信息与Transformer架构相结合,实现了对火星地表图像更精细的语义分割,在复杂地貌区域取得了较为稳定的表现。上述研究表明,深度学习方法能够在传统几何方法难以处理的复杂场景中,提供更具判别力的语义表达(Tan等,2025)。

2.2.2　泛化能力与跨域适应的深度感知方法

由于深空探测任务具有跨区域、跨任务和跨平台等特点,深度学习模型的泛化能力成为制约其工程应用的重要因素。针对不同探测任务和数据分布之间存在显著差异的问题,有研究提出通过混域训练策略(Vincent等,2022),将来自不同任务的数据

共同用于模型训练，从而提升模型在多任务场景下的适应能力。此外，为进一步增强模型对未知环境的鲁棒性，一些工作引入了约束聚类等无监督或弱监督策略，通过在特征空间中引入结构性约束，提升深度特征的判别性和一致性（Panambur 和 Parente，2025）。这类方法在一定程度上缓解了深空视觉任务中仿真到真实场景差距所带来的性能退化问题，为深度学习模型在深空环境中的工程应用提供了新的思路。

2.2.3 深空星表多模态感知方法

在深空探测任务中，视觉感知往往并非孤立目标，而是与导航安全性、地形可通行性分析等高层任务紧密相关。为此，一些研究开始探索将深度学习方法与多模态信息相结合，以提升视觉感知结果在实际任务中的可用性。Castilla-Arquillo 等人（2025）通过融合可见光、深度和热成像等多源信息，实现了对非结构化地形的多模态语义分割，在复杂光照和遮挡条件下展现出更稳定的性能。与此同时，部分研究尝试将深度语义分割结果直接引入遍历性分析和路径安全评估中，构建面向任务的混合感知架构，从而实现从像素级理解到任务级判断的感知能力提升（Chiuchiarelli 等，2024）。这类方法体现了深度学习视觉感知由单一识别任务向系统集成方向发展的趋势。

尽管深度学习方法在深空视觉感知任务中展现出良好的性能潜力，但其工程化应用仍面临一系列现实挑战。深空探测系统对可靠性、可解释性和可验证性具有严格要求，而深度神经网络通常被视为“黑箱模型”，其决策机制难以直接解释。为应对这一问题，有研究开始探索面向深空应用的可解释性工具与分析框架，以增强模型在实际任务中的可信度（Lundstrom 等，2022）。

总体来看，基于深度学习的视觉感知方法通过数据驱动方式显著提升了深空探测系统在复杂地形语义理解与环境建模方面的能力，尤其在非结构化场景下表现出相较传统方法更强的表达能力。从方法发展脉络来看，相关研究正逐步由早期依赖充分标注数据的监督学习模式，向小样本泛化以及结构约束引导的学习范式演进，其核心在于通过学习式表征弥补传统几何方法在复杂环境中的建模不足，使模型能够更好地刻画地形结构与语义关系。然而，这一类方法的性能较大程度上仍依赖于数据分布与训练策略，其在深空场景中的适用范围与稳定性仍有进一步拓展空间。

具体而言，现有研究在多个层面上仍值得进一步深入。从数据角度看，深空探测任务普遍面临样本规模有限与标注获取困难的问题，同时真实数据与仿真数据之间存在一定分布差异，这在一定程度上对模型的跨场景泛化能力提出了更高要求；从模型角度看，深度网络通过高维参数学习复杂特征表示，在提升语义表达能力的同时，其内部决策机制的可解释性仍有待进一步增强，以更好地契合深空任务对可靠性与可验证性的需求；从系统角度看，随着模型规模不断增长，其计算复杂度与资源消耗也随之增加，而深空探测器平台在算力与能耗方面具有严格约束，因此如何在性能与效率之间取得更优平衡，成为当前研究的重要关注点。进一步来看，现有方法在提升表达能力与满足工程部署需求之间仍需进行更精细的权衡，这也为后续方法设计提供了重要启示。

因此，基于深度学习的视觉感知方法已显著提升了深空探测系统对复杂地形的语义理解能力，并推动视觉感知由局部几何测量向高层环境表征不断演进。但从深空任务场景出发，这类方法仍需进一步适应长期自主运行所提出的更高要求：一方面，月球、火星等未知复杂深空环境具有非结构化、先验知识欠缺和高不确定性的特点，要求感知系统不仅能够识别地形，还需要在连续任务过程中稳定支撑导航、规划与作业等后续决策；另一方面，深空探测通常面临长时延通信、有限人工干预以及边端资源受限等现实条件，这意味着视觉感知结果不仅要具有较高精度，还应具备足够的稳定性、一致性和工程可用性。未来深空视觉感知研究需要更加关注复杂任务链路中的长期适应能力与系统支撑作用，从而为后续具身智能决策方法的发展提供更加坚实的感知基础。

2.3 深空探测感知数据集及评价指标

数据集是深空探测视觉感知方法研究的重要基础，其规模、标注形式以及任务覆盖范围在很大程度上决定了算法设计思路与性能评估方式。与地面视觉任务相比，深空探测场景下的数据获取受制于探测成本高、任务周期长以及通算资源受限等因素，公开可用的数据集数量有限，且在标注完整性和任务多样性方面存在明显不足。因此，围绕不同探测天

体构建和整理具有代表性的数据集，并形成统一、可复现的评价指标体系，是推动深空探测视觉感知方法发展的关键环节。

2.3.1 月球探测数据集

月球作为人类深空探测的重要目标之一，其探测数据集主要来源于轨道遥感影像、着陆器和巡视器获取的月球表面图像。早期月球视觉研究多依赖于公开的遥感影像资源，如中国科学院国家天文台（李春来 等，2010）以及美国国家航空航天局（NASA）发布的高分辨率月球轨道影像，这类数据为地形建模和特征分析提供了基础观测支持，但通常缺乏面向机器学习任务的系统标注。

近年来，随着月球自主探测和巡视任务的发展，一些面向智能感知任务的数据集开始出现。Liu等人（2024）提出针对月球自主探测场景集立体图像、深度信息、点云以及语义标签于一体的多传感器数据集，可支持语义分割、视觉导航和三维重建等多种任务。该数据集在一定程度上弥补了月球场景下多任务标注数据不足的问题，为复杂感知与决策方法的研究提供了统一测试基准。

此外，为缓解真实月球数据稀缺的问题，部分研究通过仿真和合成手段构建月球地表数据集。例如，LunarSeg数据集利用模拟月球环境生成包含障碍物与可通行区域标注的RGB-Depth图像（Jiao等，2025），用于验证语义分割和避障算法在月球场景下的适用性。

2.3.2 火星探测数据集

相比月球，火星探测任务开展时间更长，积累的数据资源也更为丰富，是当前深空视觉感知研究中数据集最为集中的方向之一。NASA行星数据系统长期对外开放火星轨道器和巡视器采集的原始影像数据（Swan等，2021），为火星视觉研究提供了基础数据来源（Wagstaff等，2021）。

在此基础上，Tan等人（2024）进一步构建了面向机器学习任务的标注数据集，并给出了多类别地形标签，广泛用于地形分类和语义理解研究。围绕我国“天问一号”火星探测任务，也有研究开始整理和发布针对“祝融号”巡视器图像的标注数据集（曹学欢 等，2025）。Lyu等人（2023）对火星地表岩石和障碍物进行了人工标注，可用于语义分割和目标检测算法的评估，为深空智能感知研究提供了重要支撑。此外，一些针对具体工程需求的专用数据集也相继提出，例如，面向月球和火星落石检测任务构建的岩石/落石标注数据集，为目标检测和危险识别算法提供了评测平台（Bickel等，2021）。在轨道尺度上，NOAH-H（novelty or anomaly hunter-HiRISE）数据集利用HiRISE高分辨率影像对火星地貌进行分类，验证了深度学习方法在大尺度深空遥感场景中的适用性（Barrett等，2022）。

2.3.3 其他深空天体与仿真数据集

除月球和火星外，一些研究开始关注小天体和其他深空环境的数据集构建。AstroVision数据集汇集了来自多种小天体的图像样本，并提供较为丰富的标注信息，可用于特征检测和跨天体泛化研究（Driver等，2023）。这类数据集为探索视觉感知方法在不同深空环境中的通用性提供了新的实验基础。

同时，由于真实深空数据获取成本高昂，模拟和增强型数据集在深空视觉研究中仍将长期存在。这类数据集在模型预训练、参数敏感性分析以及极端场景测试中具有不可替代的价值，但其与真实探测环境之间的域差异问题仍需在应用中谨慎对待。

2.3.4 深空探测视觉任务的评价指标

针对深空探测视觉感知任务，研究中通常采用与计算机视觉领域相似的评价指标，但需结合具体任务特点进行解读。在地形分类和语义分割任务中，分类准确率、召回率以及交并比（intersection over union，IoU）是最常用的评价指标，用于衡量模型对不同地形类别的识别能力（Li等，2024）。

在目标检测和危险识别任务中，平均精度（average precision，AP）和召回率广泛用于评估检测性能。对于面向导航和实时决策的边缘感知系统，推理时间、帧率以及计算资源消耗等工程指标同样具有重要意义，往往与感知精度共同构成综合评价体系。

总体而言，现有深空探测数据集在天体类型、任务覆盖和标注粒度等方面呈现出明显的不均衡特征。月球和火星数据集相对丰富，但仍难以全面覆盖复杂任务需求，其他深空环境的数据资源则更加有限。这种数据层面的约束在一定程度上限制了视觉感知方法，尤其是深度学习方法在深空场景中的推广应用，也进一步凸显了构建高质量、多任务深空数据集以及统一评价标准的重要性。

3　深空探测具身智能方法

3.1　深空探测具身智能任务类型

深空探测具身智能系统面向月球、火星及更远深空环境中的长期自主运行需求，其核心目标是在复杂、未知且通信受限的条件下，实现探测器对环境的持续感知、自主决策与可靠执行。从任务层面来看，深空探测具身智能并非单一能力的简单叠加，而是由一系列相互关联的自主任务构成，贯穿探测器从环境感知、路径规划到具体作业执行的全过程。根据任务功能和决策层级的不同，如表2所示，现有深空探测任务通常可以概括为自主视觉导航、自主任务规划以及自主科学操作3类典型任务类型。

3.1.1　深空自主视觉导航

自主视觉导航是深空探测具身智能系统的基础任务，其目标是使探测器在缺乏高精度外部定位和持续地面指令支持的情况下，依托自身视觉感知能力完成环境理解、位姿估计和安全移动（王大轶和黄翔宇，2009）。在月球和火星等非结构化地表环境中，探测器通常面临地形起伏大、障碍分布随机以及光照条件变化剧烈等问题，传统依赖精确地图和外部导航信号的方式难以长期适用。因此，自主视觉导航任务强调探测器基于实时获取的视觉信息，对周围环境进行持续感知与判断，从而支持路径选择、避障决策以及运动控制等导航行为。这类任务在火星巡视器长期自主行走任务中具有关键作用，也是后续高层任务规划和作业执行得以开展的前提条件。

3.1.2　深空自主任务规划

在完成基本导航能力的基础上，深空探测任务通常需要探测器在较长时间尺度内自主组织和调整任务执行顺序，即自主任务规划（徐瑞 等，2021）。该类任务关注探测目标选择、作业时序安排以及资源分配等问题，其复杂性不仅来源于环境不确定性，还与任务目标多样性和系统资源受限密切相关。

表2　深空探测具身智能方法研究归纳

Table 2　Overview of embodied intelligence methods for deep space exploration

任务类别	文献	方法类别	数据来源	主要任务
自主视觉导航	Bajracharya等人（2008）	视觉姿态估计	火星探测车	火星车自主导航
	Yu等人（2021）	端到端深度学习	月球车仿真数据	月球车路径规划
	Hu等人（2021）	深度强化学习	月球车仿真环境	月球车导航与路径规划
	Daftry等人（2022）	机器学习	火星探测车真实与仿真数据	火星车安全路径规划
	Verma等人（2023）	导航系统集成	“毅力号”多传感器数据	火星车自动驾驶
	Gao等人（2024）	深度强化学习	月球探测车	月球车路径规划
	Santra等人（2025）	跨域强化学习	月球仿真数据	跨域导航策略迁移
自主任务规划	徐瑞等人（2021）	局部自主规划	深空探测任务\系统数据	科学作业任务规划
	孙泽翼等人（2024）	多智能体强化学习	深空任务仿真数据	探测器自主任务规划
自主科学操作	Estlin等人（2012）	规则驱动目标采样	MER火星车	自动科学目标识别与观测
	Francis等人（2017）	规则驱动目标采样	“好奇号”火星车	自动科学观测任务执行
	Stefanuk和Skonieczny（2022）	自编码器异常检测	火星探测车	行星表面未知目标
	苗仕林等人（2025）	深度学习	月球空间站数据	空间站行为识别

在深空探测场景中，由于通信时延大、地面干预成本高，探测器往往需要在有限先验信息条件下，根据环境变化动态调整任务计划。例如，在火星探测任务中，巡视器需要在保证安全的前提下，平衡科学探测收益、能量消耗以及时间约束等多重目标。因此，自主任务规划被视为连接环境感知与具体执行行为的重要决策层任务。

3.1.3　深空自主科学操作

自主科学操作是深空探测具身智能系统中最贴近物理执行层的任务类型，主要涉及探测器对具体

作业行为的自主完成，如样品采集、机械臂操作、仪器部署及精细对位等。在深空环境，这类任务通常具有操作精度要求高、失败代价大以及环境反馈不确定等特点，对系统自主性提出了更高要求（孙辉先等，2017）。

由于地面遥操作在深空场景下面临显著的通信时延和带宽限制，自主作业控制逐渐成为提升探测效率和降低任务风险的重要方向。例如，在火星样品采集和原位分析任务中，探测器需要在有限感知信息支持下，自主完成复杂操作流程，从而减少对实时地面控制的依赖（朱立颖 等，2019）。这类任务不仅考验感知与控制的协同能力，也对系统整体智能水平提出了更高要求。

3.2 基于遥控遥操的深空探测具身方法

在深空探测早期阶段，受限于计算能力、传感器水平以及人工智能成熟度，探测器的运行高度依赖地面控制系统，形成了以遥控遥操为核心的传统控制模型。该模式下，探测任务规划、路径设计以及关键操作决策主要由地面控制中心完成，探测器在星体表面或轨道环境中按照预先下发的指令序列执行既定动作。这种控制范式在早期月球和火星探测任务中得到了广泛应用，并在保障任务安全性和系统可控性方面发挥了重要作用（苗仕林 等，2025）。

典型的遥控遥操模型通常采用“感知—回传—决策—下发—执行”的闭环结构，即由探测器获取环境数据并回传地面，经过人工分析与规划后生成控制指令，再通过深空通信链路发送至探测器执行。因此，系统的智能主要体现在地面控制端，而探测器本体则以执行预定义指令为主。该模式在通信条件良好、环境变化相对缓慢的任务中具有较高可靠性，例如早期的火星着陆器与巡视器任务，以及部分月球探测工程（Crisp等，2003）。

为降低地面操作风险并提升任务效率，研究者在遥控遥操框架内逐步引入了基于规则的局部自主能力。例如，通过预定义行为模式和安全约束，使探测器能够在特定条件下自主完成简单避障、姿态调整和异常处理等操作（于登云 等，2020）。这类方法通常依赖人工设计的控制逻辑和规则，具备良好的可解释性和可验证性。我国在嫦娥工程和火星探测任务中，也大量采用了此类遥控为主、规则自主为辅的控制策略，以保障复杂深空任务的可靠实施（叶培建和彭兢，2006）。

然而，随着探测任务向更复杂环境和更长任务周期发展，传统遥控遥操模型的局限性逐渐显现。首先，深空环境存在显著的通信带宽受限和计算资源受限问题，难以支持全方位、高可靠的实时遥操作（安建平 等，2016）。其次，依赖人工的控制方式难以应对非结构化环境中的突发变化，容易导致任务效率下降甚至安全风险增加。此外，基于规则的自主控制方法在面对复杂、多目标任务时缺乏灵活性，其适应能力和扩展性受到明显限制（Bajracharya等，2008）。总体来看，传统深空探测控制模型在任务安全性和系统稳定性方面具有不可替代的工程价值，是深空探测发展历程中的重要阶段。但其高度依赖地面干预和人工经验的特点，也在客观上制约了探测器自主能力的进一步提升。这些问题为后续引入基于学习和模型推理的具身智能方法奠定了现实背景和技术动机。

3.3 基于深度学习的深空探测具身方法

相较于传统遥控遥操模式，深度学习驱动的局部自主具身方法强调在不彻底改变深空任务工程体系的前提下，将学习模型嵌入探测器关键环节，实现对感知、局部决策与作业执行能力的增强。这类方法通常保留传统导航与任务管理框架，例如安全约束、规则逻辑与地面协同机制，但以学习模型替代人工特征设计或经验阈值，使探测器在局部场景理解、风险评估与操作目标识别方面具备更强的自适应能力。总体而言，深度学习在深空具身系统中的应用呈现出由感知增强向决策辅助，再到自主科学作业的逐级演进路径，为后续更高层次的全自主智能奠定了可落地的工程基础。

3.3.1 基于深度学习的自主导航与任务规划

在自主视觉导航任务中，相关研究一方面延续了传统状态估计方法，另一方面逐步引入数据驱动模块以增强地形可通行性分析能力。以火星车任务为例，高锡珍等人（2024）通过无迹卡尔曼滤波（unscented Kalman filter，UKF）利用所建立的测量模型和状态递推模型实现探测器位置、速度和姿态实时估计。与此同时，面向着陆区与行驶区的自动化可通行性分析也逐步引入数据驱动模块，用于地形分类、岩石检测与数字高程建模，以提升大尺度地形评估效率并降低人工负担（Ono等，2016）。这一阶段的核心特点是：学习模型主要用于感知增强，导航决策仍以传统规划控制为主体，因此具备较好的工

程可解释性与可控性。

深空探测任务规划具有典型的强约束与长时序特性，需要同时满足安全、高效与科学收益等多目标约束。传统任务规划高度依赖人工与规则系统，而深度学习的引入更多体现为对不确定性的适应以及对局部策略的自动化学习。在规划增强路径上，数据驱动规划首先表现为规划输入质量提升和规划评估过程加速。进一步地，深度学习旨在保留安全关键操作的前提下，利用学习模型减少搜索开销或提升路径筛选效率。典型代表如MLNav框架，强调对安全关键、资源受限的火星车规划流程进行学习增强，通过机器学习对候选路径进行更有效的排序与筛选，从而减少高代价物理约束评估的调用次数，提高规划效率并兼顾安全性（Daftry 等，2022）。类似地，Abcouwer等人（2021）提出面向火星车体系的学习型启发式代价估计方法，用于更优排序候选路径，降低规划阶段的计算负担。此类方法体现了局部自主的关键特征：学习模型不直接替代安全关键模块，而是在工程框架内承担加速与辅助角色，更易落地部署。

在更高自主化诉求下，一部分研究开始探索端到端学习导航与深度强化学习路径规划，尤其在月球与类月场景中形成了相对系统的研究链条。例如，Yu等人（2021）面向月球车提出安全约束下的端到端学习路径规划方法，通过构建加载真实月面地形数据的仿真环境，实现从感知输入到规划输出的整体学习；Hu等人（2021）进一步尝试用深度强化学习实现月球车大范围自主导航与路径规划。与此同时，跨域迁移也成为增强泛化能力的重要方向，如Santra等人（2025）探讨将地面场景训练的导航策略零样本迁移到类月环境的可行性。这类端到端范式在复杂非结构化场景中展现出更强的自适应潜力，但其不足也较为突出：高度依赖仿真训练与数据分布，且在安全可证性、故障可控性以及极端场景鲁棒性方面仍难以满足严苛的深空工程要求。同时，强化学习逐步被用于学习面向月球车等平台的规划策略。Gao等人（2024）提出面向月球车状态空间的深度强化学习规划方法，显示出在复杂问题上较传统规划器更高的求解效率与速度优势。

总体来看，深度学习驱动的自主导航与任务规划在深空探测中的优势主要体现在利用数据驱动方式提升规划效率，并在不确定环境中提供更具适应性的策略候选。但是，目前方法仍有进一步完善空间，例如其训练过程较依赖仿真环境与数据分布，在硬约束满足与安全验证方面尚缺乏统一框架，且任务规划往往涉及跨子系统协同，如何融入现有任务管理体系仍需系统工程层面的闭环验证。

3.3.2　基于深度学习的自主科学操作

自主科学操作的核心在于探测器能够在尽可能非人为干预情况下主动识别科学价值目标，并优先回传关键数据，从而提升单位时间的科学产出。早期代表性成果是OASIS/AEGIS系列工作（Estlin等，2012），其以机器学习和图像分析为基础，实现对岩石等地质目标的自动识别与仪器自主瞄准，使巡视器在执行行驶任务的同时能够自主发现并采集高价值科学观测数据。随着系统成熟，AEGIS用于“好奇号”火星车ChemCam化学相机的自主目标选择并进入常态化科学运行阶段，显著增强了机会性科学观测能力与观测效率（Francis等，2017）。这一阶段的关键价值在于：在不改变总体任务链路的前提下，通过“局部自主、地面审核”的方式提升科学作业的效率与覆盖范围，体现出典型的半智能范式特征。

近年来，深度学习进一步推动自主科学操作从二分类扩展到任务驱动的目标选择。因此，面向火星类环境的岩石检测与分割研究持续增多，例如Furlán等人（2019）利用卷积神经网络（convolutional neural network，CNN）实现类火星环境下的岩石检测与分割，为导航与作业提供目标级语义信息；Kuang等人（2021）提出面向巡视器导航视觉的岩石分割网络框架，有助于提升复杂地形下的环境建模能力。与此同时，面向星上在轨处理的深度学习流水线也开始系统化，Marek和Nalepa（2024）提出面向行星车星上处理的端到端深度学习流程，将岩石检测、语义分割与资源约束纳入统一考虑，体现出深空边缘侧部署需求驱动下的方法演进趋势。

此外，围绕稀缺数据与未知目标的自主科学需求，深度生成模型方法被用于异常检测与下游数据筛选。Stefanuk和Skonieczny（2022）针对“好奇号”火星车数据提出面向行星表面图像的新颖性检测方法，引入变分自编码器等模型以提升异常发现能力，为有限带宽下的数据优先回传提供了关键支撑。更进一步，随着“毅力号”任务推进，基于自主机器人技术的科学运行能力持续增强，相关研究总结了“毅力号”在火星表面通过自主系统执行科学目标采集的

实践经验，体现出深度学习与自治系统在工程任务中的可验证价值。

综合而言，基于深度学习的自主科学操作方法能够显著提升科学目标发现与观测效率，并在通信受限条件下优化数据下行策略；与此同时，这类方法在复杂光照、域偏移与工程可靠性等方面仍有进一步优化空间，例如，模型对域偏移与光照变化敏感、误检与漏检可能带来科学资源浪费，且科学操作链路对可靠性与可解释性要求极高，深度模型如何在工程任务中实现稳定运行仍是关键挑战。

3.4 基于多模态大模型的深空探测具身方法

随着多模态大模型与视觉—语言模型在跨模态理解、复杂推理和任务规划方面能力的显著提升，研究者开始探索将多模态大模型引入深空探测具身系统，构建面向“全局自主智能”的新范式。不同于传统遥控遥操依赖地面人员闭环决策，也区别于深度学习阶段以局部感知增强为主的半智能体系，多模态大模型更强调在统一的语义空间内实现“感知—规划—执行”的贯通，能够以自然语言为媒介对多源观测、任务目标与执行约束进行显式表达与联合推理，从而形成更高层次的自主任务生成与决策编排能力（Driess等，2023）。从发展阶段看，该方向在深空领域仍处于探索早期，但其潜在能力边界正在被逐步验证，并为下一代深空探测自主体系提供新的技术支点。

如图2所示，在具身智能通用研究中，多模态大模型已形成较清晰的方法谱系：一类是分层式多模态语言模型，通过将图像、状态量等连续观测嵌入到语言模型的词向量序列中，使模型具备跨模态的联合推理能力，从而在单一模型内同时完成场景问答、任务分解与策略生成等功能；另一类是端到端视觉—语言—动作（vision-language-action，VLA）模型，直接将动作离散化为可预测的词向量，使模型能够从视觉观测与语言指令出发输出可执行动作序列（Zitkovich等，2023）。进一步地，基于视频生成与动力学技术的世界—动作模型提供了显式推理新方法，即由大模型推理出未来任务的规划与步骤执行结果，依据逆向动力学推理出可执行动作序列，从而在保持可控性的同时提升系统对复杂任务的表达与组合能力（Brohan等，2023）。这些通用具身框架为深空探测提供了可复用的系统范式：上层以大模型承担任务规划，下层以安全可证的控制模块完成执行闭环。

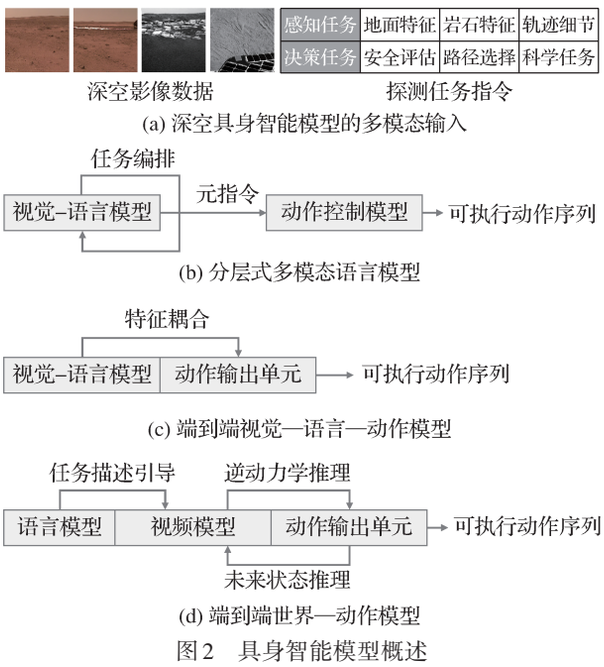


图2 具身智能模型概述

Fig. 2 Overview of embodied intelligence model

((a) multimodal inputs for deep-space embodied intelligence model; (b) hierarchical multimodal language model; (c) end-to-end vision-language-action model; (d) end-to-end world-action model)

面向航天任务本体，例如探测器、卫星也已出现一些代表性工作。例如，Rodriguez-Fernandez等人（2024）提出语言模型作为航天器操作者的研究设想，讨论将大模型作为指导、导航与控制决策过程中的代理模块，使其能够根据用户指令和系统状态输出可执行操作序列，从而提升自主运行与决策效率。Maranto（2024）进一步提出卫星大模型框架，将大模型作为航天器的高层推理与控制中枢，在仿真任务中评估其目标导向决策能力，并指出当任务复杂度提升时，单纯依赖大模型推理会出现性能退化。这类研究表明，大模型更适合作为高层认知与决策模块接入深空系统。

在行星巡视器场景中，多模态大模型同样展现出对复杂地形理解与策略选择的潜力。由于月球、火星地表具备强非结构化特征，仅依赖几何或低层视觉特征往往难以稳定刻画地形语义与风险等级，而视觉—语言模型能够通过自然语言形成更可解释的地形理解中间层，为导航策略选择提供语义支撑。Cheng等人（2025）提出多模态大模型驱动的多模式系统，将视觉—语言模型用于解析地形复杂度并选

择不同导航模式，从而在安全与效率之间进行自适应折中。这类混合架构代表了当前深空领域较现实的落地路线：以大模型增强高层理解与决策，将安全关键部分保留在可验证的规划控制链路中。

综合来看，多模态大模型驱动的全自主智能体系在深空探测中展现出重要潜力。首先，这类模型能够在统一语义空间中处理视觉观测、状态信息与任务指令，实现多源信息的联合建模与跨模态推理，从而为复杂任务提供具备自然语言解释能力的中间表达和可组合的任务分解机制；其次，借助“规划—工具调用—执行反馈”的代理式闭环，大模型有望将任务执行由静态指令响应进一步扩展到动态重规划与多步推理过程，使系统在开放环境下具备更强的任务适应能力；同时，在深空通信受限、人工干预成本高的背景下，这类方法使部分高层决策能够在星上完成，从而减少对地面实时介入的依赖，契合未来深空任务向高自主、低干预方向发展的总体需求（Beard，2024）。从场景角度看，这种能力的意义并不局限于单次决策性能提升，而在于增强探测器面对长时序任务链路时的任务组织能力，使系统在巡视、探测、采样与科学观测等不同阶段之间形成更加连贯的语义衔接与策略延续。

进一步来看，多模态大模型在深空场景中的工程价值主要体现在对复杂任务理解、决策编排和流程组织能力的增强，而非对底层控制链路的直接替代。由于深空任务通常伴随长时间跨度、强环境不确定性以及严格的安全约束，模型在极端场景下的鲁棒决策能力仍有待进一步增强，其与底层控制模块之间的协同方式也需要在后续研究中持续完善。尤其是在深空数据稀缺、真实环境与仿真环境存在分布差异的条件下，这类方法的稳定应用更加依赖系统级验证与分层式架构设计。这样的体系结构既顺应了深空探测对更高自主性的需求，也更符合现阶段深空任务对可靠性与工程可控性的总体要求。

3.5 深空探测具身数据集及评价指标

与面向分类、检测和分割等静态视觉任务的数据集不同，深空探测具身智能数据集更强调时序性、交互性与闭环性，即数据不仅包含视觉观测，还需同步提供位姿、控制指令、资源状态以及任务回报等信息，以支持导航、任务规划与自主科学操作的端到端评测与对比。由于真实深空任务数据获取成本极高且标注困难，现阶段具身数据主要以高保真仿真基准为核心载体，通过多源传感器同步、真实物理渲染以及可复现实验流程，为算法验证提供可控的评测环境。因此，深空具身智能数据集的构建目标并非单纯扩大样本规模，而在于尽可能覆盖复杂地形、极端光照与资源受限等典型任务条件，从而支撑具身系统在非结构化深空场景中的可靠性评估（李德仁等，2025）。

在导航与定位方面，高保真仿真数据集正逐步从单任务数据走向多模态导航闭环基准。面向月面、火星表面自主导航的具身数据集，通常需要同时支持“环境感知—定位建图—路径跟踪”的闭环评估，因此在数据形态上更强调多传感器同步、轨迹真值以及可复现实验场景。近年来，一类代表性工作是基于仿真或数字孪生构建的导航基准数据集，例如 LuSNAR（lunar segmentation，navigation and reconstruction dataset）数据集通过 Unreal Engine 构建多场景月面环境，提供立体相机、激光雷达和惯性测量单元等多源观测及位姿真值，可用于语义分割、三维重建与导航等任务的联合评测（Liu 等，2024）。此外，邢琰等人（2022）提出了“地外场景感知—操控—状态闭环”的多源数据集天衢 1.0，为地外探测智能感知与操控研究提供了数据基础与验证支撑，推动深空探测人工智能理论方法与技术应用研究。

在自主科学操作与机会性探测方面，具身智能数据集的核心目标是提升科学产出并降低地面干预依赖。AEGIS（autonomous exploration for gathering increased science）系统是该方向的代表性工程实践之一，其能够在火星车上自主选择科学目标并触发后续观测，支持在不额外增加地面通信与人工分析成本的条件下提高科学观测效率（Verma 等，2023）。此外，Li 等人（2025a）围绕火星等深空探测中典型的非结构化复杂场景，构建了面向长时序任务的具身智能数据集，为复杂地形下无人系统的感知、决策与操作研究提供了重要数据支撑。在此基础上，作者进一步提出了分层式智能无人系统视觉—语言—动作模型，为深空探测器提供高层决策规划与底层执行器自主操作能力，实现了复杂指令在动态环境中的稳定完成，为探火工程中火星车在非结构化场景中自主运行提供前期验证依据。

针对上述具身任务数据形态，评价指标体系通常从“安全性—有效性—效率—鲁棒性”4 个维度展开：在自主导航中，常用指标包括导航成功率、轨迹

长度、能量消耗、危险区域侵入率以及定位误差与地图一致性等。在任务规划中，评测重点转向任务完成率、硬约束满足率、重规划次数以及在线计算开销等，以量化系统在不确定环境中的自适应能力与资源管理能力。在自主科学操作中，常用指标包括目标发现的命中率与有效观测比例、单位时间科学产出、无效观测率、回传数据压缩与优先级排序收益等，以衡量自治系统对科学回报的提升幅度。近年来也出现将任务级可信验证纳入评测框架的趋势，例如通过任务失败模式统计、系统级安全约束违规率，推动具身智能走向可证可靠。

总体而言，深空具身智能数据集与评测仍处于快速演进阶段，高保真仿真数据与挑战基准正在加速形成，为导航、规划和科学操作提供可复现的系统级对比平台。面向未来，深空具身智能评测体系需要进一步走向统一化、闭环化和高保真化，并通过跨任务通用指标刻画系统级自主能力，以安全认证与可信评测流程支撑深空任务的工程落地与长期自主、高效、可靠运行。

4 结 语

本文围绕深空探测这一典型复杂未知的应用场景，系统梳理了具身智能相关方法的研究进展。从任务特性与工程约束出发，本文分析了深空具身智能面临的核心挑战，并围绕视觉感知方法、具身智能决策模型以及数据集与评价体系，对国内外相关研究进行了归纳与总结。通过对传统方法、深度学习方法与多模态大模型方法的系统比较，可以看出，深空探测具身智能正由面向单一功能的局部优化，逐步演进为面向长期自主作业的系统闭环构建。

总结而言，深空探测具身智能在场景理解、任务规划与自主科学作业等方面已取得初步进展。然而，面向长期、自主、可靠的深空探测任务需求，现有方法距离形成可持续的系统化能力仍存在明显不足。其根本原因在于，深空任务所要求的并非受控实验条件下局部性能指标的优化，而是在未知环境、资源约束严格的综合条件下，实现感知、决策与执行链路的稳定闭环运行。由此可见，深空具身智能的发展难以简单沿用通用具身智能的既有技术范式，而需要面向深空任务特性，在环境建模、系统架构与验证范式等层面形成更具针对性的研究框架。结合本文的系统梳理，未来值得重点关注的研究方向主要体现在以下几个方面。

1）面向深空非结构化场景的感知泛化能力。深空感知的核心挑战，来自复杂环境与分布变化的共同作用。月球、火星等场景通常具有地形起伏大、结构非规则以及光照变化剧烈等特点，容易导致观测信息不稳定。同时，深空任务还面临目标稀疏、先验不足和样本有限等问题，不同区域、不同任务乃至不同载荷条件下的数据分布也往往存在明显差异。传统方法具有较好的物理可解释性，但在复杂场景中的高层语义建模能力相对有限。深度学习方法虽然提升了特征表达能力，但其性能通常依赖训练数据所覆盖的场景范围，当观测条件或任务背景发生变化时，模型鲁棒性仍需进一步增强。因此，未来研究的重点不宜局限于封闭数据集上的识别精度提升，而应更加关注开放环境中的稳定表征与持续适应能力。面向小样本、弱标注甚至无标注条件，构建兼具语义一致性与跨模态约束能力的感知模型，使其在先验不足和观测退化条件下仍能稳定识别地形结构及关键目标，将是值得持续推进的方向。相应地，数据集与评价体系也需要由静态识别性能评估，逐步拓展到对跨场景迁移能力、长期适应能力以及任务支撑能力的系统刻画。

2）面向长期自主作业的决策可信性与闭环可靠性。深空具身智能中的决策问题，本质是面向长时序任务链的可信闭环控制问题。现有方法中，传统遥控遥操模式具有较强可控性，但自主性难以满足深空探测长远发展需求，深度学习与大模型驱动的方法尽管一定程度上提升了自主决策能力，但目前的方法引入了工程偏差累积以及异常场景下策略失稳等新问题。因此，未来研究的关键在于具备工程可落地、行为可约束以及风险可控制的自主决策能力。从技术路径上看，深空自主决策可靠性的提升，需要构建面向任务约束的分层决策机制，即由高层模块承担目标感知及理解、中层模块承担约束一致的规划推理以及底层模块承担控制执行，从而在系统层面实现智能性与可控性的统一。进一步地，可信验证也不应仅依赖离线实验精度，而应围绕任务成功率、决策边界稳定性与实际部署可靠性，建立工程可信的评价框架。

3）深空具身智能的系统协同问题。从现有研究发展来看，深空具身智能正处于由单模块优化向系

统级协同演进的关键阶段。早期研究多聚焦于感知、导航、规划和操作等局部能力提升，而多模态大模型与具身智能方法的发展，则进一步推动系统具备跨模态理解与高层认知能力。相较于系统局部性能提升，深空场景更核心的问题在于系统中各层级能力的协同一致性、运行稳定性与工程可验证性。基于此，未来研究应更加关注面向任务闭环的层级协同架构构建：以多模态感知模块提供环境状态的泛化感知，以高层智能模块完成自主决策与任务分解，以中低层规划控制模块实现运动控制与安全执行，并通过记忆与反馈机制形成全链路闭环。总结来说，系统协同的关键不在于模型规模扩展或模块堆叠，而在于跨层级表示对齐、任务信息有效传递与闭环反馈机制完善，从而支撑深空自主作业系统的整体智能演进。

4）边端部署的系统轻量化问题。面向深空探测任务的实际部署场景，资源约束是相关方法实现工程落地的前提条件。受算力、存储、功耗及通信带宽等因素限制，深空平台难以长期支撑高复杂度模型的边缘处理，因此，边端部署问题的关键在于有限资源预算下感知、规划与控制等多环节协同。现有研究主要围绕模型压缩、参数裁剪与量化加速等局部技术路径展开，这为复杂模型的边端适配提供了重要基础，但面向长时序、高可靠、强耦合的深空任务链路，单点层面的复杂度优化仍需进一步向系统层面拓展。未来研究需要更加关注时效性约束下性能与效率的均衡优化问题。推动边端部署优化由局部轻量化，逐步走向算法系统与硬件协同设计，从而更好地支撑深空具身智能系统的长期、稳定与可靠在轨运行。

综上，深空探测具身智能是在极端非结构化地外环境、长时序任务与高可靠约束共同作用下形成的一类新型系统智能问题。其未来突破的关键，不仅在于单一技术的持续提升，更在于围绕感知泛化、决策可信、系统协同与资源约束适配等核心问题，建立面向深空任务特点的统一技术框架与验证范式。本文期望通过对相关研究脉络、方法体系与关键问题的系统梳理，为该方向后续研究提供更清晰的分析视角，并为未来深空探测任务中更高自主性与更高可靠性的智能系统构建提供参考。

参考文献（References）

Abcouwer N, Daftry S, del Sesto T, Toupet O, Ono M, Venkatraman S, et al. 2021. Machine learning based path planning for improved rover navigation//Proceedings of 2021 IEEE Aerospace Conference. Big Sky, MT, USA: IEEE: 1-9 [DOI: 10.1109/AERO50100.2021.9438337]

An J P, Jin S, Xu J, Zhang Y and Shao L W. 2016. Development and outlook of deep space communication network protocol. Journal on Communications, 37(7): 50-61 (安建平, 靳松, 许军, 张宇, 邵立伟. 2016. 深空通信网络协议的发展与展望. 通信学报, 37(7): 50-61) [DOI: 10.11959/j.issn.1000-436x.2016133]

Bajracharya M, Maimone M W and Helmick D. 2008. Autonomy for Mars rovers: past, present, and future. Computer, 41(12): 44-50 [DOI: 10.1109/MC.2008.479]

Barrett A M, Balme M R, Woods M, Karachalios S, Petrocelli D, Joudrier L, et al. 2022. NOAH-H, a deep-learning, terrain classification system for Mars: results for the ExoMars Rover candidate landing sites. Icarus, 371: #114701 [DOI: 10.1016/j.icarus.2021.114701]

Beard B L. 2024. Concept of operations for increasingly autonomous space operations. Journal of Space Safety Engineering, 11(4): 636-651 [DOI: 10.1016/j.jsse.2024.06.008]

Bickel V T, Mandrake L and Doran G. 2021. A labeled image dataset for deep learning-driven rockfall detection on the Moon and Mars. Frontiers in Remote Sensing, 2: #640034 [DOI: 10.3389/frsen.2021.640034]

Brohan A, Chebotar Y, Finn C, Hausman K, Herzog A, Ho D, et al. 2023. Do as I can, not as I say: grounding language in robotic affordances//Proceedings of the 6th Conference on Robot Learning. Auckland, New Zealand: PMLR: 287-318

Cao X H, Peng M, Wan W H, Wang B, Wang Y X, Di K C, et al. 2025. Convolutional self-attention network for automated rock detection in complex Martian terrain. National Remote Sensing Bulletin, 29(10): 3020-3033 (曹学欢, 彭嫚, 万文辉, 王彪, 王晔昕, 邸凯昌, 等. 2025. 火星表面石块的卷积自注意力网络识别方法. 遥感学报, 29(10): 3020-3033) [DOI: 10.11834/jrs.20255015]

Castilla-Arquillo R, Perez-del-Pulgar C, Gerdes L, Garcia-Cerezo A and Olivares-Mendez M A. 2025. OmniUnet: a multimodal network for unstructured terrain segmentation on planetary rovers using RGB, depth, and thermal imagery [EB/OL]. [2025-08-01]. https://arxiv.org/pdf/2508.00580.pdf

Cheng S N, Zhou R Y, Feng W H, Yang H G, Gao H B, Deng Z Q, et al. 2025. VLM-empowered multi-mode system for efficient and safe planetary navigation//Proceedings of 2025 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Hang-

zhou, China: IEEE: 1487-1494 [DOI: 10.1109/IROS60139.2025.11247594]

Chiuchiarelli A, Franchini G, Messina F and Chiaberge M. 2024. Towards safer planetary exploration: a hybrid architecture for terrain traversability analysis in Mars rovers [EB/OL]. [2026-03-08]. https://arxiv.org/pdf/2410.17738.pdf

Crisp J A, Adler M, Matijevic J R, Squyres S W, Arvidson R E and Kass D M. 2003. Mars exploration rover mission. Journal of Geophysical Research: Planets, 108(E12): #8061 [DOI: 10.1029/2002JE002038]

Daftry S, Abcouwer N, Sesto T D, Venkatraman S, Song J L, Igel L, et al. 2022. MLNav: learning to safely navigate on Martian terrains. IEEE Robotics and Automation Letters, 7(2): 5461-5468 [DOI: 10.1109/LRA.2022.3156654]

Driess D, Xia F, Sajjadi M S M, Lynch C, Chowdhery A, Ichter B, et al. 2023. PaLM-E: an embodied multimodal language model// Proceedings of the 40th International Conference on Machine Learning. Honolulu, USA: PMLR: #340

Driver T, Skinner K A, Dor M and Tsiotras P. 2023. AstroVision: towards autonomous feature detection and description for missions to small bodies using deep learning. Acta Astronautica, 210: 393-410

Estlin T A, Bornstein B J, Gaines D M, Anderson R C, Thompson D R, Burl M, et al. 2012. AEGIS automated science targeting for the MER opportunity rover. ACM Transactions on Intelligent Systems and Technology (TIST), 3(3): #50 [DOI: 10.1145/2168752.2168764]

Francis R, Estlin T, Doran G, Johnstone S, Gaines D, Verma V, et al. 2017. AEGIS autonomous targeting for ChemCam on Mars Science Laboratory: deployment and results of initial science team use. Science Robotics, 2(7): #4582 [DOI: 10.1126/scirobotics.aan4582]

Fu X J, Cai X D, Liu Y M, Tian G L and Wen X. 2023. Comprehensive testing technology for multi spacecraft collaboration in Chang'E-7 lunar probe. Journal of Deep Space Exploration, 10(6): 577-585 (傅晓晶, 蔡晓东, 刘一鸣, 田国亮, 温新. 2023. "嫦娥七号"探测器多器协同综合测试技术. 深空探测学报, 10(6): 577-585) [DOI: 10.15982/j.issn.2096-9287.2023.20230103]

Furlán F, Rubio E, Sossa H and Ponce V. 2019. Rock detection in a Mars-like environment using a CNN//Proceedings of the 11th Mexican Conference on Pattern Recognition. Querétaro, Mexico: Springer: 149-158 [DOI: 10.1007/978-3-030-21077-9_14]

Gao A, Lu S Y, Xu R, Li Z Y, Wang B, Zhu S Y, et al. 2024. Deep reinforcement learning based planning method in state space for lunar rovers. Engineering Applications of Artificial Intelligence, 127: #107287 [DOI: 10.1016/j.engappai.2023.107287]

Gao X Z, Huang X Y and Xu C. 2024. Intelligent fusion autonomous navigation method for Mars precise landing. Journal of Deep Space Exploration, 11(1): 24-30 (高锡珍, 黄翔宇, 徐超. 2024. 火星精确着陆智能融合自主导航方法. 深空探测学报, 11(1): 24-30) [DOI: 10.15982/j.issn.2096-9287.2024.20230041]

Geng X and Xu Q. 2022. A fast photogrammetric processing method for linear pushbroom planetary remote sensing images. Journal of Deep Space Exploration, 9(3): 348-356 (耿迅, 徐青. 2022. 一种线阵推扫式行星遥感影像摄影测量快速处理方法. 深空探测学报, 9(3): 348-356) [DOI: 10.15982/j.issn.2096-9287.2022.20210002]

Goh E, Chen J D and Wilson B. 2022. Mars terrain segmentation with less labels//Proceedings of 2022 IEEE Aerospace Conference (AERO). Big Sky, USA: IEEE: 1-10 [DOI: 10.1109/AERO53065.2022.9843245]

Hu T, Cao T, Zheng B, Zhang H M and Ni M Y. 2021. Large-scale autonomous navigation and path planning of lunar rover via deep reinforcement learning//Proceedings of 2021 China Automation Congress (CAC). Beijing, China: IEEE: 2050-2055 [DOI: 10.1109/CAC53003.2021.9728075]

Jiao S F, Zeng Z W, Su Z Q, Chen X Y L, Zhou Z T and Lu H M. 2025. LuSeg: efficient negative and positive obstacles segmentation via contrast-driven multi-modal feature fusion on the lunar//Proceedings of 2025 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Hangzhou, China: IEEE: 5962-5969 [DOI: 10.1109/IROS60139.2025.11246916]

Jin W M, Chen J G, Run D, Cheng Y H and Zhao X Q. 2024. Mission planning technology for flight control of Mars exploration. Journal of Deep Space Exploration, 11(5): 485-495 (金文马, 陈俊刚, 润冬, 程艳合, 赵兴乾. 2024. 火星探测飞行控制任务规划技术研究. 深空探测学报, 11(5): 485-495) [DOI: 10.15982/j.issn.2096-9287.2024.20220107]

Kuang B Y, Wisniewski M, Rana Z A and Zhao Y F. 2021. Rock segmentation in the navigation vision of the planetary rovers. Mathematics, 9(23): #3048 [DOI: 10.3390/math9233048]

Li C L, Liu J J, Ren X, Mou L L, Zou Y L, Zhang H B, et al. 2010. The global image of the Moon obtained by the Chang'E-1: data processing and lunar cartography. Science China Earth Sciences, 53(8): 1091-1102 (李春来, 刘建军, 任鑫, 牟伶俐, 邹永廖, 张洪波, 等. 2010. 嫦娥一号图像数据处理与全月球影像制图. 中国科学: 地球科学), 40(3): 294-306 [DOI: 10.1360/zd2010-40-3-294]

Li D R, Wang M, Xiao J, Li M, Di K C, Li X, et al. 2025. On ubiquitous spatio-temporal intelligence. National Remote Sensing Bulletin, 29(6): 1388-1398 (李德仁, 王密, 肖晶, 李明, 邸凯昌, 李熙, 等. 2025. 论无所不在的时空智能. 遥感学报, 29(6): 1388-1398) [DOI: 10.11834/jrs.20255016]

Li D X, He S B, Tian J Y, Zhang Y S, Xie W Y, Cao M X, et al. 2025a. Uni-sight: an E2E vision-language-action system unifying multi-view alignment and multi-modal fusion//Proceedings of the 33rd ACM International Conference on Multimedia. Dublin, Ireland: ACM: 7142-7151 [DOI: 10.1145/3746027.3754495]

Li D X, Xie W Y, Wang Z X, Lu Y B, Li Y S and Fang L Y. 2024. FedDiff: diffusion model driven federated learning for multi-modal and

multi-clients. IEEE Transactions on Circuits and Systems for Video Technology, 34(10): 10353-10367 [DOI: 10.1109/TCSVT.2024.3407131]

Li D X, Zhang Y S, Cao M X, Liu D L, Xie W Y, Hui T L, et al. 2025b. Towards long-horizon vision-language-action system: reasoning, acting and memory//Proceedings of 2025 IEEE/CVF International Conference on Computer Vision. Honolulu, USA: IEEE: 6839-6848 [DOI: 10.1109/ICCV51701.2025.00643]

Li L C, Li J, Wang J, Yu T Y, Liu C K, Li N, et al. 2022. Design and construction of a simulating and testing system for the remote operation of planetary surface exploration. Journal of Deep Space Exploration, 9(1): 72-79 (李立春, 李剑, 王镓, 于天一, 刘传凯, 李楠, 等. 2022. 天体表面探测遥操作模拟验证系统设计与实现. 深空探测学报, 9(1): 72-79) [DOI: 10.15982/j.issn.2096-9287.2022.20210133]

Li L H, Lian J, Guo L and Wang R B. 2013. Visual odometry for planetary exploration rovers in sandy terrains. International Journal of Advanced Robotic Systems, 10(5): #234 [DOI: 10.5772/56342]

Li M D, Zhang X H, Ding Y L, Fan X Y, Zhang B, Xu C, et al. 2025. Research progress on autonomous navigation technology for landing and roaming of extraterrestrial surface detectors. Journal of Deep Space Exploration, 12(2): 110-123 (李茂登, 张兴华, 丁运来, 范项媛, 张兵, 徐超, 等. 2025. 地外星表探测器着陆—巡视自主导航技术研究进展. 深空探测学报, 12(2): 110-123) [DOI: 10.15982/j.issn.2096-9287.2025.20240085]

Liu J Y, Zhang Q Y, Wan X, Zhang S Y, Tian Y L, Han H D, et al. 2024. LuSNAR: a lunar segmentation, navigation and reconstruction dataset based on muti-sensor for autonomous exploration [EB/OL]. [2026-03-08]. https://arxiv.org/pdf/2407.06512.pdf

Liu J Z, Ge P, Kang Y, Zhang T X, Jiang Y C, Ma K, et al. 2024. Study on the construction scheme of mega interconnected knowledge systems in deep space exploration. Journal of Deep Space Exploration, 11(1): 79-89 (刘继忠, 葛平, 康焱, 张天馨, 姜亦宸, 马克, 等. 2024. 深空探测科学目标谱系构建方案探索. 深空探测学报, 11(1): 79-89) [DOI: 10.15982/j.issn.2096-9287.2024.20230169]

Lundstrom D, Huyen A, Mevada A, Yun K and Lu T. 2022. Explainability tools enabling deep learning in future in-situ real-time planetary explorations//Proceedings of 2022 IEEE Aerospace Conference (AERO). Big Sky, USA: IEEE: 1-8 [DOI: 10.1109/AERO53065.2022.9843288]

Lyu W K, Wei L H, Zheng D, Liu Y and Wang Y M. 2023. MarsNet: automated rock segmentation with transformers for Tianwen-1 mission. IEEE Geoscience and Remote Sensing Letters, 20: #3506605 [DOI: 10.1109/LGRS.2022.3227338]

Ma Y, Li Z J, Wu B and Duan R. 2025. DepthFormer: depth-enhanced transformer network for semantic segmentation of the Martian surface from rover images. Earth and Space Science, 12(6): #e2024EA003812 [DOI: 10.1029/2024EA003812]

Ma Y Q, Liu S C, Jia Y, Jia Y H, Zhang J L, Wei S Y, et al. 2014. Experimental research of navigation and localization algorithm based on stereo images for the lunar rover. Scientia Sinica Technologica, 44(10): 1097-1104 (马友青, 刘少创, 贾阳, 贾永红, 张建利, 魏士俨, 等. 2014. 基于立体图像的月球车导航定位试验研究. 中国科学: 技术科学), 44(10): 1097-1104 [DOI: 10.1360/n092013-00031]

Maimone M, Cheng Y and Matthies L. 2007. Two years of visual odometry on the Mars exploration rovers. Journal of Field Robotics, 24(3): 169-186 [DOI: 10.1002/rob.20184]

Maranto D. 2024. LLMSat: a large language model-based goal-oriented agent for autonomous space exploration [EB/OL]. [2026-03-08]. https://arxiv.org/pdf/2405.01392.pdf

Marek D and Nalepa J. 2024. End-to-end deep learning pipeline for onboard extraterrestrial rock segmentation. Engineering Applications of Artificial Intelligence, 127: #107311 [DOI: 10.1016/j.engappai.2023.107311]

Miao S L, Tang Y P, Zhang C J, Liu C K, Zhang J T and Cao J F. 2025. Intelligent temporal behavior detection for lunar space station. Journal of Image and Graphics, 30(9): 3039-3049 (苗仕林, 唐业鹏, 张淳杰, 刘传凯, 张济韬, 曹建峰. 2025. 面向月球空间站的智能时序行为检测. 中国图象图形学报, 30(9): 3039-3049) [DOI: 10.11834/jig.250135]

Mohammad F, Gao Y, Kay S, Field R, De Benedetti M and Ntagiou E V. 2024. Deep learning based semantic segmentation for Mars rover terrain classification//Proceedings of 2024 International Conference on Space Robotics (iSpaRo). Luxembourg, Luxembourg: IEEE: 292-298 [DOI: 10.1109/iSpaRo60631.2024.10687827]

Ning X L, Yang Y Q, Fang J C and Wu W R. 2023. The progress of autonomous celestial navigation for deep space spacecraft. Journal of Deep Space Exploration, 10(2): 99-108 (宁晓琳, 杨雨青, 房建成, 吴伟仁. 2023. 深空探测器自主天文导航研究进展. 深空探测学报, 10(2): 99-108) [DOI: 10.15982/j.issn.2096-9287.2023.20230023]

Olson C F, Matthies L H, Wright J R, Li R X and Di K C. 2007. Visual terrain mapping for Mars exploration. Computer Vision and Image Understanding, 105(1): 73-85 [DOI: 10.1016/j.cviu.2006.08.005]

Ono M, Rothrock B, Almeida E, Ansar A, Otero R, Huertas A, et al. 2016. Data-driven surface traversability analysis for Mars 2020 landing site selection//Proceedings of 2016 IEEE Aerospace Conference. Big Sky, USA: IEEE: 1-12 [DOI: 10.1109/AERO.2016.7500597]

Panambur T and Parente M. 2025. Enhancing Martian terrain recognition with deep constrained clustering. IEEE Transactions on Geoscience and Remote Sensing, 63: #4600515 [DOI: 10.1109/TGRS.2025.3561687]

Rodriguez-Fernandez V, Carrasco A, Cheng J, Scharf E, Siew P M and Linares R. 2024. Language models are spacecraft operators [EB/

OL]. [2026-03-08]. https://arxiv.org/pdf/2404.00413.pdf

Rothrock B, Kennedy R, Cunningham C, Papon J, Heverly M and Ono M. 2016. SPOC: deep learning-based terrain classification for Mars rover missions//Proceedings of the AIAA SPACE 2016. Long Beach, USA: AIAA: #5539 [DOI: 10.2514/6.2016-5539]

Santra S, Robbins T and Yoshida K. 2025. Transferable deep reinforcement learning for cross-domain navigation: from farmland to the Moon [EB/OL]. [2026-03-08]. https://arxiv.org/pdf/2510.23329.pdf

Stefanuk B and Skonieczny K. 2022. Novelty detection in rover-based planetary surface images using autoencoders. Frontiers in Robotics and AI, 9: #974397 [DOI: 10.3389/frobt.2022.974397]

Sun H X, Li H J, Zhang B M, Zhou C Y, Xue C B, Zhu Y, et al. 2017. Achievements and prospect of payloads technology in Chinese lunar and deep space exploration. Journal of Deep Space Exploration, 4(6): 495-509 (孙辉先, 李慧军, 张宝明, 周昌义, 薛长斌, 朱岩, 等. 2017. 中国月球与深空探测有效载荷技术的成就与展望. 深空探测学报, 4(6): 495-509) [DOI: 10.15982/j.issn.2095-7777.2017.06.001]

Sun Z Y, Wang B, Hu X Y, Xiong X and Jin H P. 2024. Multi-agent reinforcement learning autonomous task planning for deep space probes. Journal of Deep Space Exploration, 11(3): 244-255 (孙泽翼, 王彬, 胡馨月, 熊新, 金怀平. 2024. 深空探测器多智能体强化学习自主任务规划. 深空探测学报, 11(3): 244-255) [DOI: 10.15982/j.issn.2096-9287.2024.20230159]

Swan R M, Atha D, Leopold H A, Gildner M, Oij S, Chiu C, et al. 2021. AI4MARS: a dataset for terrain-aware autonomous driving on Mars//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Nashville, USA: IEEE: 1982-1991 [DOI: 10.1109/CVPRW53098.2021.00226]

Tan X M, Xi B B, Xu H T, Li J J, Li Y S, Xue C B, et al. 2024. A lightweight framework with knowledge distillation for zero-shot Mars scene classification. IEEE Transactions on Geoscience and Remote Sensing, 62: #4601516 [DOI: 10.1109/TGRS.2024.3470526]

Tan X M, Xi B B, Xue C B, Li Y S, Zheng T and Gamba P. 2025. A Mamba-based zero-shot Mars scene classification framework//Proceedings of 2025 IEEE International Geoscience and Remote Sensing Symposium. Brisbane, Australia: IEEE: 4989-4993 [DOI: 10.1109/IGARSS55030.2025.11243340]

Tang Y H, Liang Q C, Lei Z Y, Jiang T T, Xu Y, Yang Z H, et al. 2025. Status and prospect of space-borne communication technology for deep space. Journal of Deep Space Exploration, 12(5): 463-475 (唐玉华, 梁启超, 雷振羽, 姜天天, 徐杨, 杨志华, 等. 2025. 深空星载通信技术发展现状与展望. 深空探测学报, 12(5): 463-475) [DOI: 10.3724/j.issn.2096-9287.2025.20250064]

Verma V, Maimone M W, Gaines D M, Francis R, Estlin T A, Kuhn S R, et al. 2023. Autonomous robotics is driving Perseverance rover's progress on Mars. Science Robotics, 8(80): #eadi3099 [DOI: 10.1126/scirobotics.adi3099]

Vincent G, Yepremyan A, Chen J D and Goh E. 2022. Mixed-domain training improves multi-mission terrain segmentation//Proceedings of 2022 European Conference on Computer Vision. Tel Aviv, Israel: Springer Nature Switzerland: 96-111 [DOI: 10.1007/978-3-031-25056-9_7]

Wagstaff K, Lu S, Dunkel E, Grimes K, Zhao B D, Cai J, et al. 2021. Mars image content classification: three years of NASA deployment and recent advances//Proceedings of the 35th AAAI Conference on Artificial Intelligence. [s.l.]: AAAI Press: 15204-15213 [DOI: 10.1609/aaai.v35i17.17784]

Wang D Y and Huang X Y. 2009. Survey of autonomous navigation and control for deep space exploration. Aerospace Control and Application, 35(3): 6-12, 43 (王大轶, 黄翔宇. 2009. 深空探测自主导航与控制技术综述. 空间控制技术与应用, 35(3): 6-12), 43) [DOI: 10.3969/j.issn.1674-1579.2009.03.002]

Whetsel C, Levine J S, Hoffman S J, Luckey C M, Watts K D and Antonsen E L. 2025. Utilizing Martian samples for future planetary exploration—characterizing hazards and resources. Proceedings of the National Academy of Sciences of the United States of America, 122(2): #e2404251121 [DOI: 10.1073/pnas.2404251121]

Xing Y, Jiang T T, Jia Y, Liu L and Yang M F. 2022. Multisource dataset for intelligent precise perception and manipulation of extraterrestrial object surface. Aerospace Control and Application, 48(4): 9-16 (邢琰, 姜甜甜, 贾永, 刘磊, 杨孟飞. 2022. 面向地外天体表面智能感知与操控研究的多源数据集. 空间控制技术与应用, 48(4): 9-16) [DOI: 10.3969/j.issn.1674-1579.2022.04.002]

Xing Y, Teng B Y, Huang H, Gao M X, Gong J, Yu C W, et al. 2024. The technology of autonomous intelligent miniature robot and its application on lunar surface. Aerospace Control and Application, 50(6): 64-72 (邢琰, 滕宝毅, 黄煌, 高明星, 龚健, 余成武, 等. 2024. 自主智能微小机器人技术及其月面应用. 空间控制技术与应用, 50(6): 64-72) [DOI: 10.3969/j.issn.1674-1579.2024.06.008]

Xu R, Li Z Y, Zhu S Y, Wang B, Liang Z X and Shang H B. 2021. Research progress of autonomous planning technology for deep space probes. Journal of Deep Space Exploration, 8(2): 111-123 (徐瑞, 李朝玉, 朱圣英, 王棒, 梁子璇, 尚海滨. 2021. 深空探测器自主规划技术研究进展. 深空探测学报, 8(2): 111-123) [DOI: 10.15982/j.issn.2096-9287.2021.20210039]

Ye P J, Meng L Z, Ma J N, Wang Q, Li Y, Du Y, et al. 2019. Suggestions on artificial intelligence technology application and development in deep space exploration. Journal of Deep Space Exploration, 6(4): 303-316 (叶培建, 孟林智, 马继楠, 王强, 李莹, 杜宇, 等. 2019. 深空探测人工智能技术应用及发展建议. 深空探测学报, 6(4): 303-316) [DOI: 10.15982/j.issn.2095-7777.2019.04.001]

Ye P J and Peng J. 2006. Deep space exploration and its prospect in

China. Engineering Science, 8(10): 13-18 (叶培建, 彭兢. 2006. 深空探测与我国深空探测展望. 中国工程科学, 8(10): 13-18) [DOI: 10.3969/j.issn.1009-1742.2006.10.003]

Yu D Y, Zhang Z, Pan B F, Liu C K, Ding L, Zhu J H, et al. 2020. Development and trend of artificial intelligent in deep space exploration. Journal of Deep Space Exploration, 7(1): 11-23 (于登云, 张哲, 泮斌峰, 刘传凯, 丁亮, 朱继宏, 等. 2020. 深空探测人工智能技术研究与展望. 深空探测学报, 7(1): 11-23) [DOI: 10.15982/j.issn.2095-7777.2020.20190916001]

Yu X Q, Wang P and Zhang Z X. 2021. Learning-based end-to-end path planning for lunar rovers with safety constraints. Sensors, 21(3): #796 [DOI: 10.3390/s21030796]

Zhang Z, Qin T, Shi Y S, Qiao D, Jian K K, Chen H, et al. 2022. Research on artificial intelligence technology for lunar scientific research station. Journal of Deep Space Exploration, 9(6): 560-570 (张哲, 秦同, 师一帅, 乔栋, 简抗抗, 陈辉, 等. 2022. 月球科研站人工智能技术研究. 深空探测学报, 9(6): 560-570) [DOI: 10.15982/j.issn.2096-9287.2022.20220099]

Zhu L Y, Ye Z L, Li Y Q, Fu Z L and Xu Y. 2019. Modeling of autonomous flight mission intelligent planning for small body exploration. Journal of Deep Space Exploration, 6(5): 463-469 (朱立颖, 叶志玲, 李玉庆, 付中梁, 徐勇. 2019. 小天体探测自主绕飞智能规划建模. 深空探测学报, 6(5): 463-469) [DOI: 10.15982/j.issn.2095-7777.2019.05.007]

Zitkovich B, Yu T H, Xu S C, Xu P, Xiao T, Xia F, et al. 2023. RT-2: vision-language-action models transfer web knowledge to robotic control//Proceedings of the 7th Conference on Robot Learning. Atlanta, USA: PMLR: 2165-2183

作者简介

李岱勋，男，博士研究生，主要研究方向为遥感分布式在轨处理和具身智能。E-mail: ldx@stu.xidian.edu.cn

谢卫莹，通信作者，女，教授，主要研究方向为遥感图像在轨实时处理、分布式边缘计算及具身智能。

E-mail: wyxie@xidian.edu.cn

李云松，男，教授，主要研究方向为图像压缩编码和图像处理。E-mail: ysli@mail.xidian.edu.cn